

Usability Evaluation of NLP-PIER: A Clinical Document Search Engine for Researchers

Gretchen Hultman^a, Reed McEwan^b, Serguei Pakhomov^c,
Elizabeth Lindemann^d, Steven Skube^d, Genevieve B. Melton^{a,d}

^a Institute for Health Informatics, University of Minnesota, Minneapolis, Minnesota, USA

^b Academic Health Center – Information Systems, University of Minnesota, Minneapolis, Minnesota, USA

^c College of Pharmacy, University of Minnesota, Minneapolis, Minnesota, USA

^d Department of Surgery, University of Minnesota, Minneapolis, Minnesota, USA

Abstract

NLP-PIER (Natural Language Processing – Patient Information Extraction for Research) is a self-service platform with a search engine for clinical researchers to perform natural language processing (NLP) queries using clinical notes. We conducted user-centered testing of NLP-PIER's usability to inform future design decisions. Quantitative and qualitative data were analyzed. Our findings will be used to improve the usability of NLP-PIER.

Keywords:

Information Storage and Retrieval; Natural Language Processing; Evaluation Studies.

Introduction

NLP-PIER was created to provide an accessible solution through a search interface to clinical researchers interested in or requiring access to clinical NLP capabilities of clinical documents [1]. The system has two interfaces: a full text search interface and a custom interface for searching Unified Medical Language System (UMLS) concepts [2]. We were interested in understanding potential design opportunities and user acceptance of NLP-PIER and to more broadly understand the needs of clinical researchers when using a self-service NLP tool.

Methods

This study was conducted at University of Minnesota (UMN) as part of its broader clinical and translational science research platform. We designed standard tasks to use NLP-PIER and asked clinical researcher participants (n=11) to complete these tasks, usability instruments (system usability scale (SUS) [3] and NASA-Task Load Index (NASA-TLX) survey [4]), a brief interview, and exit questionnaire. Time on task, task completion percentage, and survey results were assessed. Interviews were transcribed and coded for themes

Results

For the full text search interface, questionnaire scores were 69.4 (19.8) (SUS) and 18.8 (5.7) (NASA TLX) and for the Concept search interface scores were 66.1 (32.4) and 21.8 (7.7). Average time on task and task completion varied widely. In interviews, all participants expressed that NLP-PIER was easy to use and would be useful in their work.

Conclusion

End user testing of NLP-PIER identified a number of usability challenges and several solutions. Our study also demonstrated that substantial variation exists between different users. Overall, our findings illustrate the importance of incorporating user testing and feedback in the design process.

Acknowledgements

This research was supported by the Agency for Healthcare Research & Quality (#R01HS022085 (GM)) and National Institutes of Health (#R01LM011364 (GM), #R01GM102282 (SP), #8UL1TR000114 (Blazar)).

References

- [1] R. McEwan, G.B. Melton, B.C. Knoll, Y. Wang, G. Hultman, J.L. Dale, T. Meyer, S.V. Pakhomov, NLP-PIER: a scalable natural language processing, indexing, and searching architecture for clinical notes, In: *AMIA Summits on Translational Science Proceedings*, 2016 (2016):150.
- [2] US National Library of Medicine. Unified Medical Language System (UMLS), Available from: <https://www.nlm.nih.gov/research/umls/html>.
- [3] J. Brooke, SUS-A quick and dirty usability scale, *Usability Eval Indust* **189** (1996):4-7.
- [4] S.G. Hart, NASA-task load index (NASA-TLX); 20 years later. In: *Proceedings of The Human Factors and Ergonomics Society Annual Meeting*, 2006, pp. 904-908.

Address for correspondence

Genevieve B. Melton, MD, PhD
Institute for Health Informatics and Department of Surgery
420 Delaware Street SE, Mayo Mail Code 450
Minneapolis, MN 55455
Tel: + 612-625-7992
E-mail: gmelton@umn.edu